

ПОСТРОЕНИЕ РАЗДЕЛЯЮЩЕЙ ПОВЕРХНОСТИ ДВУХ ТОЧЕЧНЫХ МНОЖЕСТВ НА ОСНОВЕ ПОИСКА БЛИЖАЙШИХ СОСЕДЕЙ

докторант, канд. техн. наук, доц. А. А. Дашкевич, НТУ "ХПИ", г. Харьков

В задачах анализа данных, машинного обучения, компьютерной графики и компьютерного зрения данные могут быть представлены в виде точечных множеств с различной геометрической структурой. При этом возникают задачи выявления свойств таких множеств, что ведёт к необходимости разработки эффективных алгоритмов на таких множествах при больших размерностях данных.

В работе предлагается применение алгоритма пространственного хеширования для вычисления точек поверхности, разделяющей два точечных множества. Получены результаты для двумерных и многомерных данных и проведено исследование полученных гиперповерхностей с точки зрения решения задачи классификации точечных множеств.

Разработанный алгоритм сводится к следующим ключевым этапам:

1. Вычисление хеш-функции для заданных точечных множеств, при этом для каждой точки также учитывается класс, к которому она принадлежит.
2. Определение тех хеш-ячеек, в которых находятся точки одновременно из двух классов.
3. Нахождение соседних с ячейками шага (2) ячеек.
4. Для всех ячеек с шагами (2) и (3) производится нахождение попарных ближайших расстояний между точками двух классов.
5. В каждом из попарных отрезков с шага (4) производится нахождение средней точки, через которую проводится кусочно-линейная гиперплоскость.
6. Нахождение пересечений соседних гиперплоскостей и уточнение разделяющей поверхности.

Предложенный метод даёт возможность определения разделяющей поверхности для двух точечных множеств, что позволяет в дальнейшем решать такие задачи, как классификация и распознавание образов, кластеризация данных, аппроксимация точечных множеств, построение выпуклых оболочек.